



# Capítulo 2 - Fundamentos

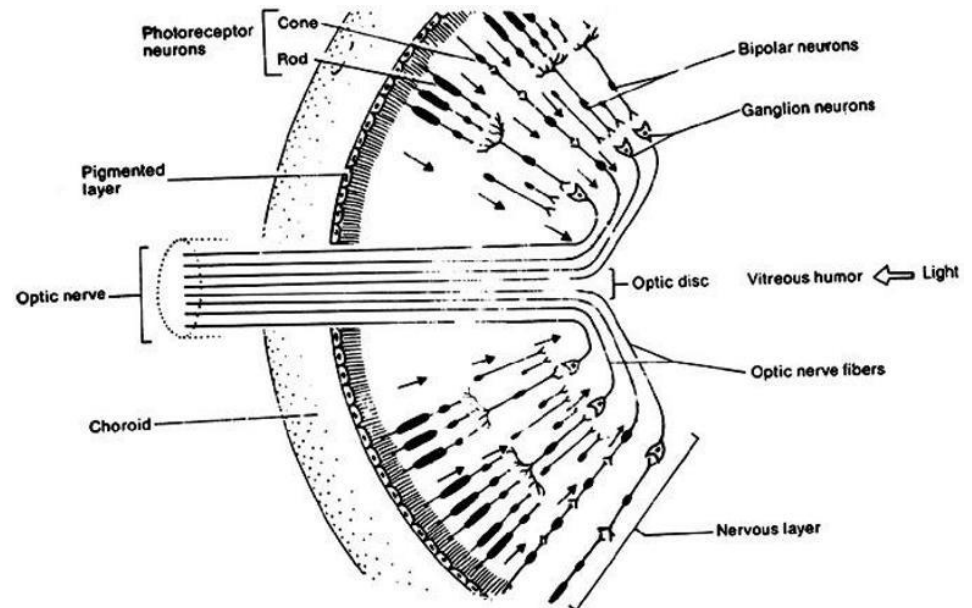
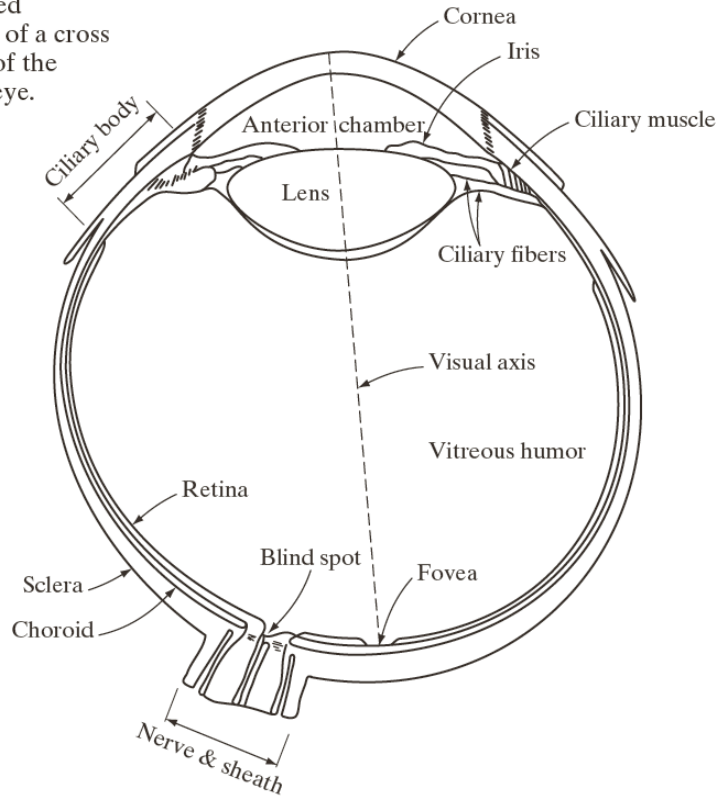
Processamento de Imagem e Visão

# Estrutura do Olho Humano (1)

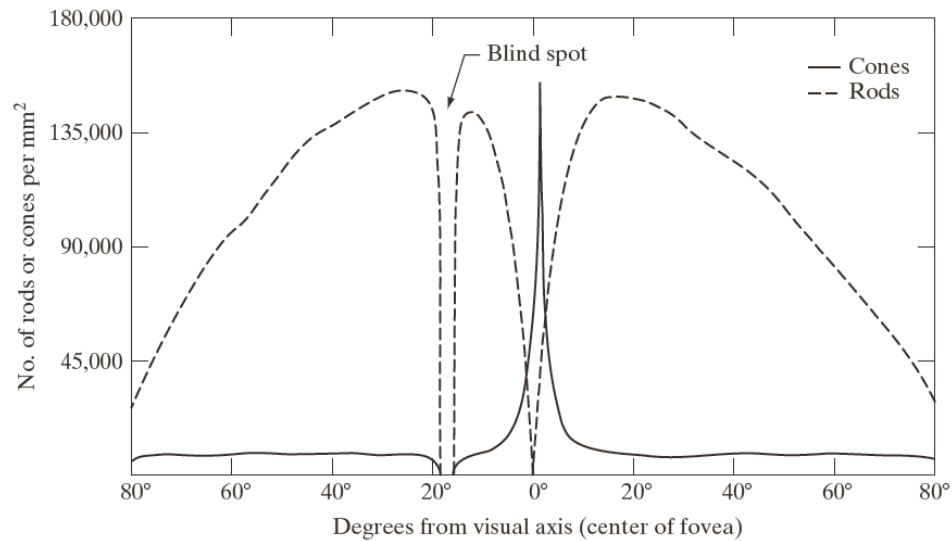


**FIGURE 2.1**

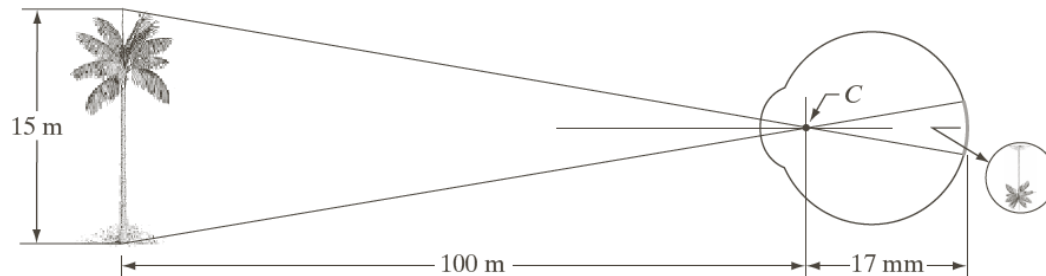
Simplified diagram of a cross section of the human eye.



# Estrutura do Olho Humano (2)



**FIGURE 2.2**  
Distribution of rods and cones in the retina.

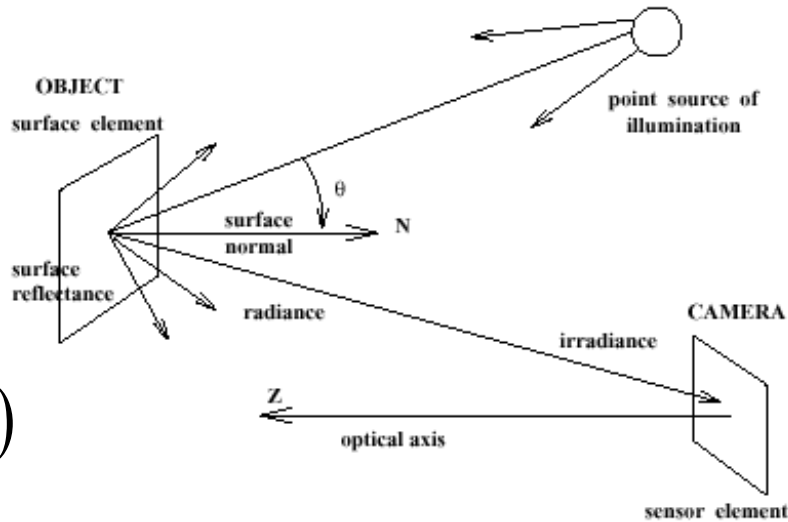


**FIGURE 2.3**  
Graphical representation of the eye looking at a palm tree. Point C is the optical center of the lens.

# Modelo fotométrico



- fonte de iluminação
- elemento de superfície
- sensor (câmara)



$$f(x, y) = i(x, y)r(x, y)$$

$0 < i(x, y) < \infty$       Quantidade de iluminação que incide no objecto

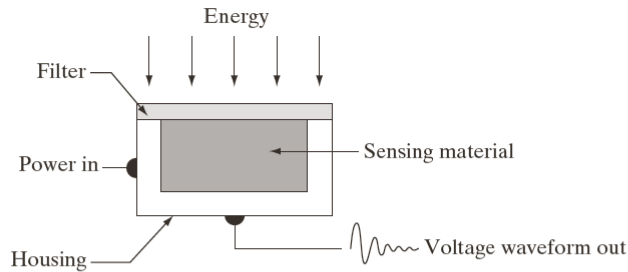
$0 < r(x, y) < 1$       Quantidade de iluminação reflectida pelo objecto

$l = f(x_0, y_0)$       onde  $L_{MIN} \leq l \leq L_{MAX}$

$[L_{MIN}, L_{MAX}]$       Intervalo de variação de intensidade luminosa

$[0, L-1]$       Normalização, onde 0 corresponde ao preto e L-1 ao branco

# Sistemas Artificiais de Aquisição de Imagens (1)



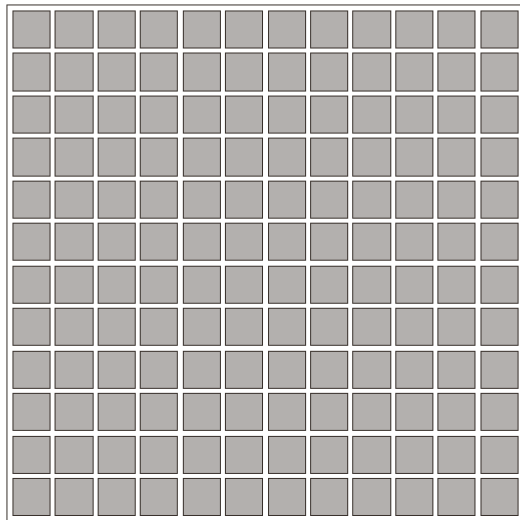
a  
b  
c

**FIGURE 2.12**

(a) Single imaging sensor.

(b) Line sensor.

(c) Array sensor.



## Câmaras CCD (*Charge-coupled devices*)

Dimensões típicas: 6.4mm x 4.8mm – ½”  
(640x480 ou 512x512 pixéis)

Para cada nova imagem, as cargas eléctricas são limpas e depois integra-se a luz recebida durante um dado intervalo de tempo (controlado pelo *shutter*)

No final, o plano de imagem (*array 2D*) é lido, varrendo cada uma das linhas

# Sistemas Artificiais de Aquisição de Imagens (2)

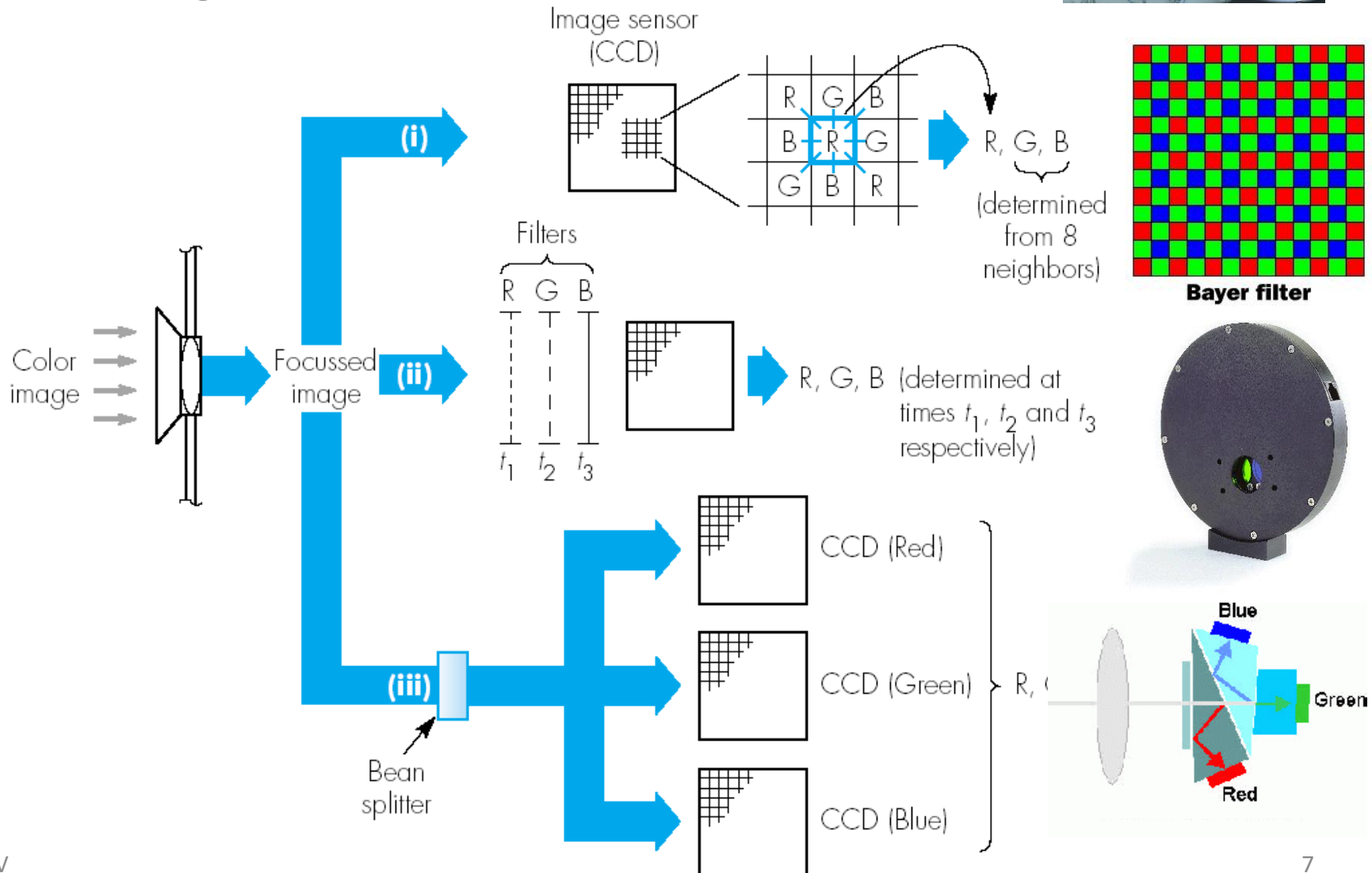


## CCD vs CMOS

- Tecnologia amadurecida;
- Tecnologia específica;
- Elevados custos de produção;
- Consumo elevado;
- Maior sensibilidade;
- Leitura sequência;
- Tecnologia recente;
- Tecnologia standard IC;
- Menos dispendiosa;
- Menor consumo;
- Menor sensibilidade;
- Amplificação por pixel;
- Acesso aleatório dos pixéis;
- Integração de outros componentes no mesmo chip;

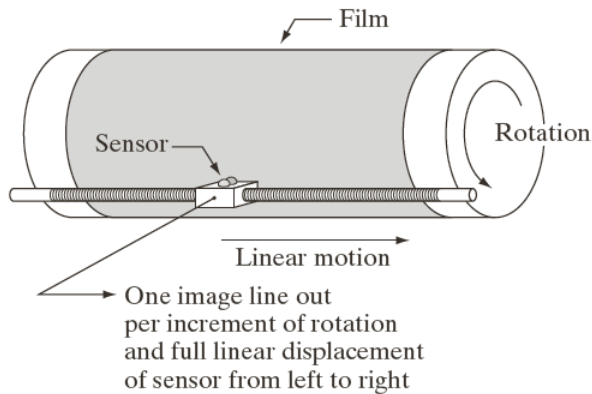


# Sistemas Artificiais de Aquisição de Imagens (3)

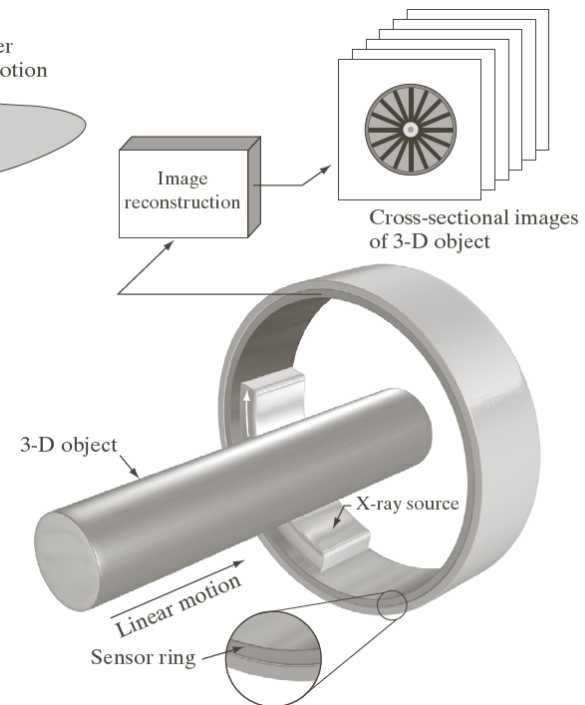
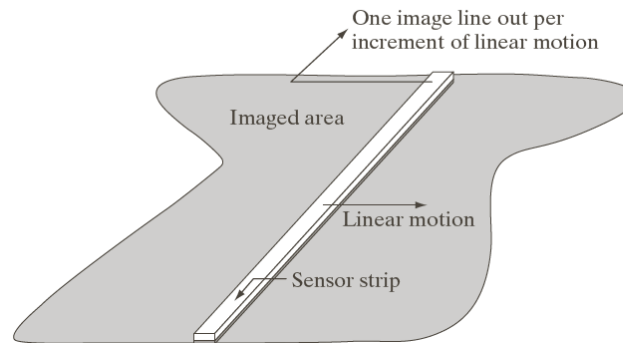




# Sistemas Artificiais de Aquisição de Imagens (4)



**FIGURE 2.13**  
Combining a single sensor with motion to generate a 2-D image.

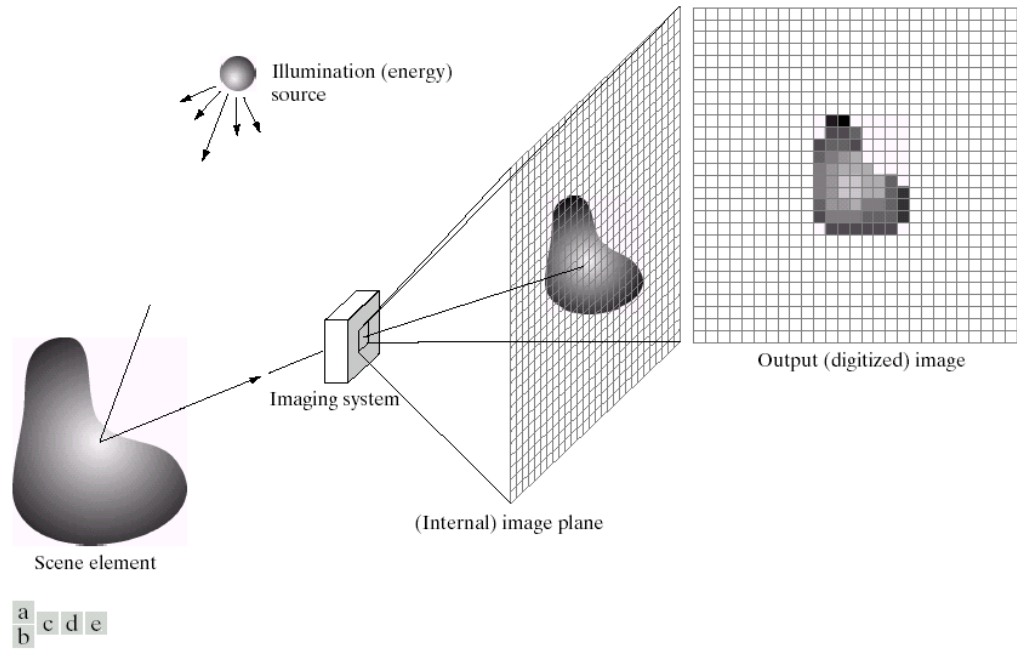


a b

**FIGURE 2.14** (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.



# Sistemas Artificiais de Aquisição de Imagens (5)

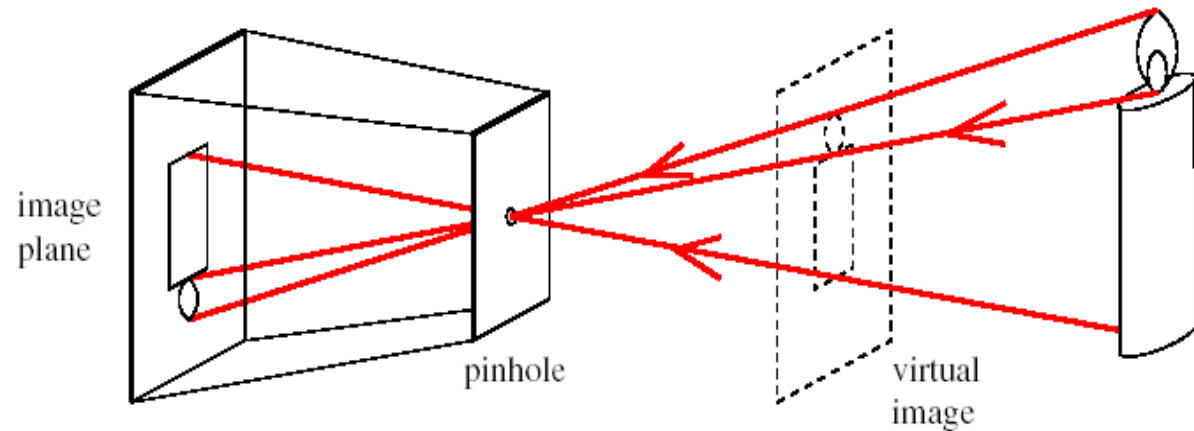


**FIGURE 2.15** An example of the digital image acquisition process. (a) Energy (“illumination”) source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

# Formação de imagem



Modelo *pin-hole* – projecção de perspectiva



Câmara escura (século XVI)

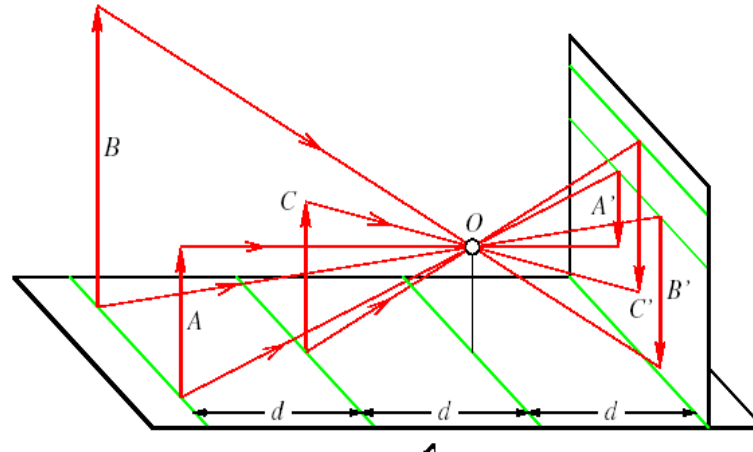
Material translúcido



# Efeitos de perspectiva



O tamanho aparente de um objecto depende da sua distância

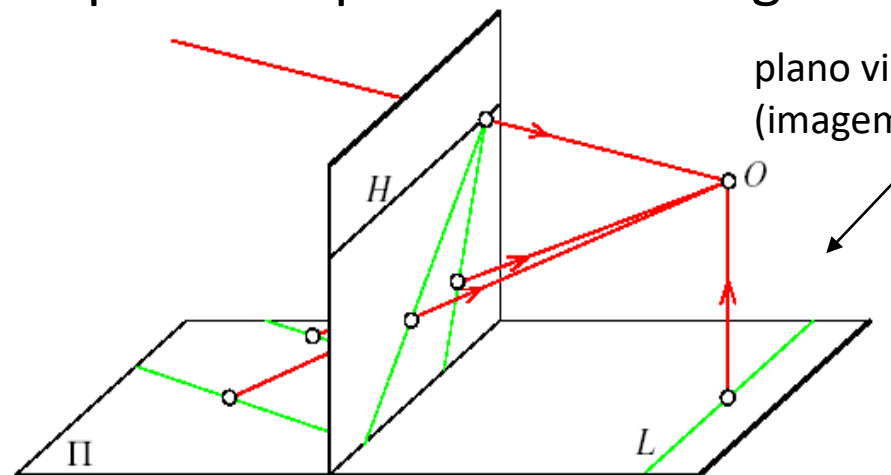


← plano da retina  
(imagem invertida)

Projecção de duas rectas paralelas parecem convergir num ponto

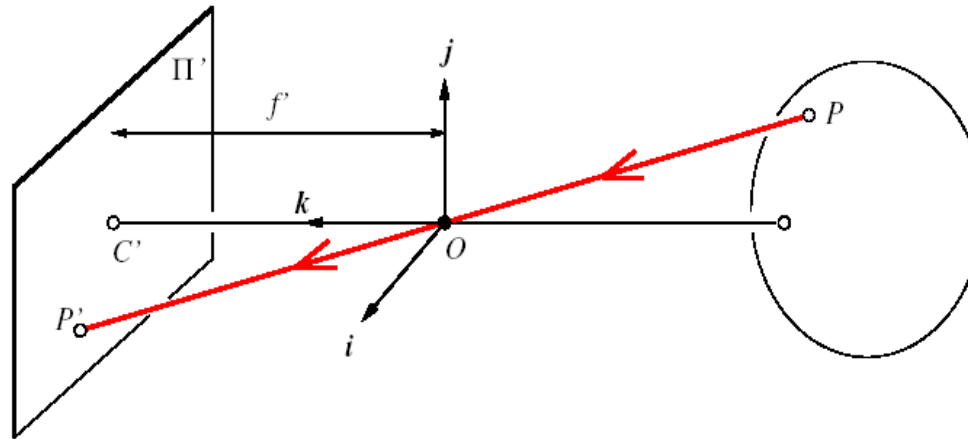


PIV



plano virtual  
(imagem não invertida)

# Projecção de Perspectiva – Modelo Matemático



$$\begin{cases} x' = -f' \frac{x}{z} \\ y' = -f' \frac{y}{z} \end{cases}$$

Como se relacionam as coordenadas de um ponto no mundo real,  $P$ , com as respectivas coordenadas na retina,  $P'$ ?

$$P = (x, y, z) \quad P' = (x', y', z')$$

Constatações:

• Os pontos  $P$ ,  $P'$  e  $O$  (centro óptico) são colineares  $\overrightarrow{OP'} = \lambda \overrightarrow{OP}$

• O ponto é projectado no plano do sensor que está colocado a uma distância  $f'$  (distância focal) do centro óptico

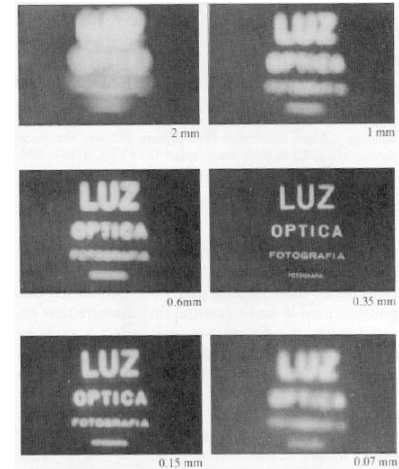
$$z' = f'$$

# Câmaras equipadas com lentes (1)

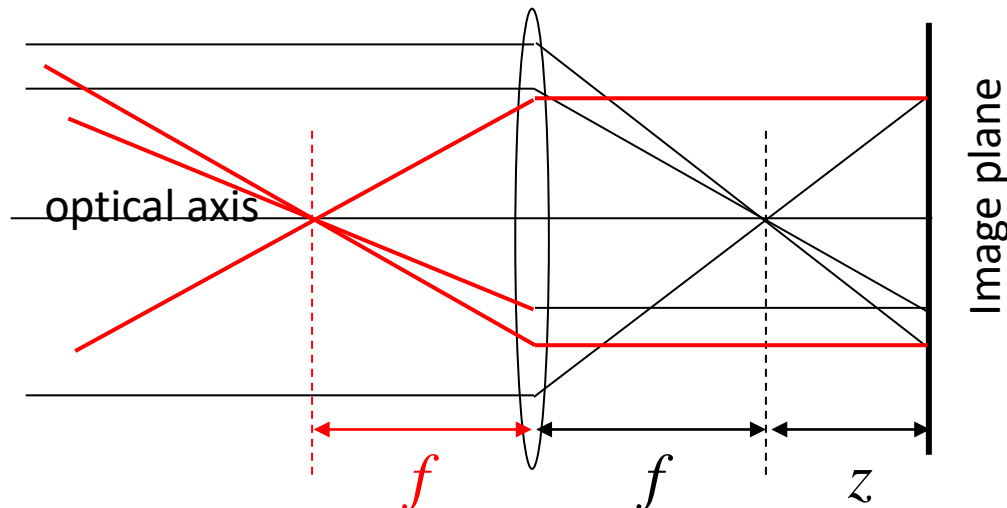


Na prática, é necessário utilizar lentes

- Lente actua como colector de luz
- Permite ajustar a focagem dos objectos (alteração de  $f$ )



fenómenos de difracção →

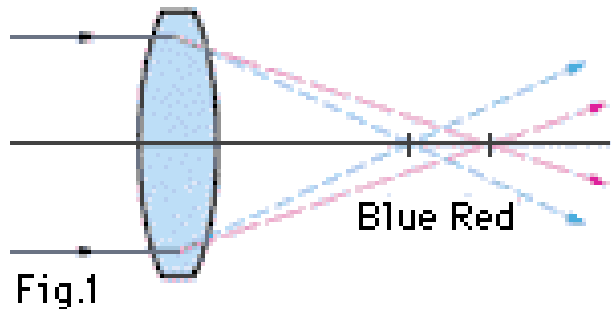


# Câmaras equipadas com lentes (2)



## Desvios relativamente ao modelo

- Imperfeições na lente conduzem a círculo de confusão
- Sensor constituído por unidades discretas; integração espacial conduz a efeito de *blurring*; limitação do detalhe observado.

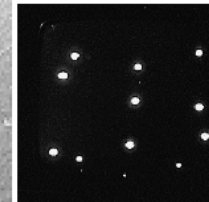


Distorção cromática

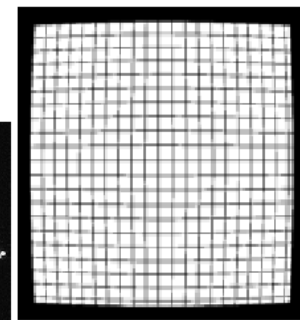
Outros problemas



*wrap-around*



*blooming*



Distorção geométrica

# Distância focal da lente (1)

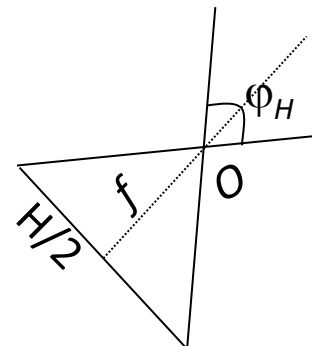
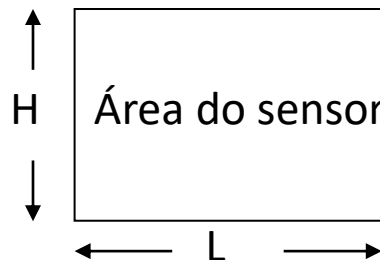


Factor de ampliação

$$m = -\frac{f'}{z} \qquad \frac{|\Delta x'|}{|\Delta x|} = \frac{|\Delta y'|}{|\Delta y|} = m$$

Campo de visão (*field of view*) – espaço da cena projectado no sensor

- Depende não só da distância focal,  $f$ , mas também da dimensão do sensor (usualmente 1/4", 1/3" ou 1/2")
- Quando  $\varphi$  elevado diz-se que é uma lente grande angular
- Quando  $\varphi$  baixo diz-se que é uma lente telescópica



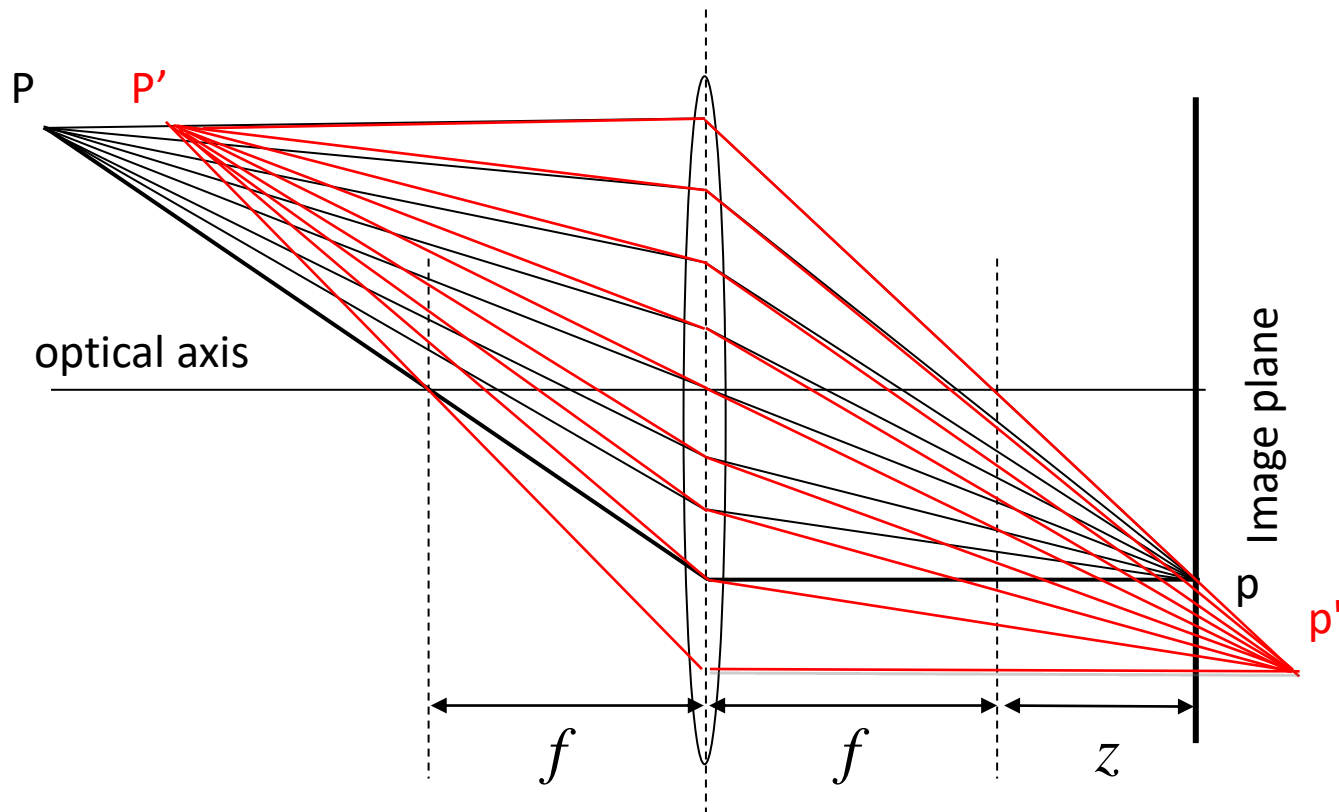
$$\varphi_H = 2 \tan^{-1} \frac{H}{2f}$$

$$\varphi_L = 2 \tan^{-1} \frac{L}{2f}$$



# Distância focal da lente (2)

Profundidade de Campo (*depth-of-field*)



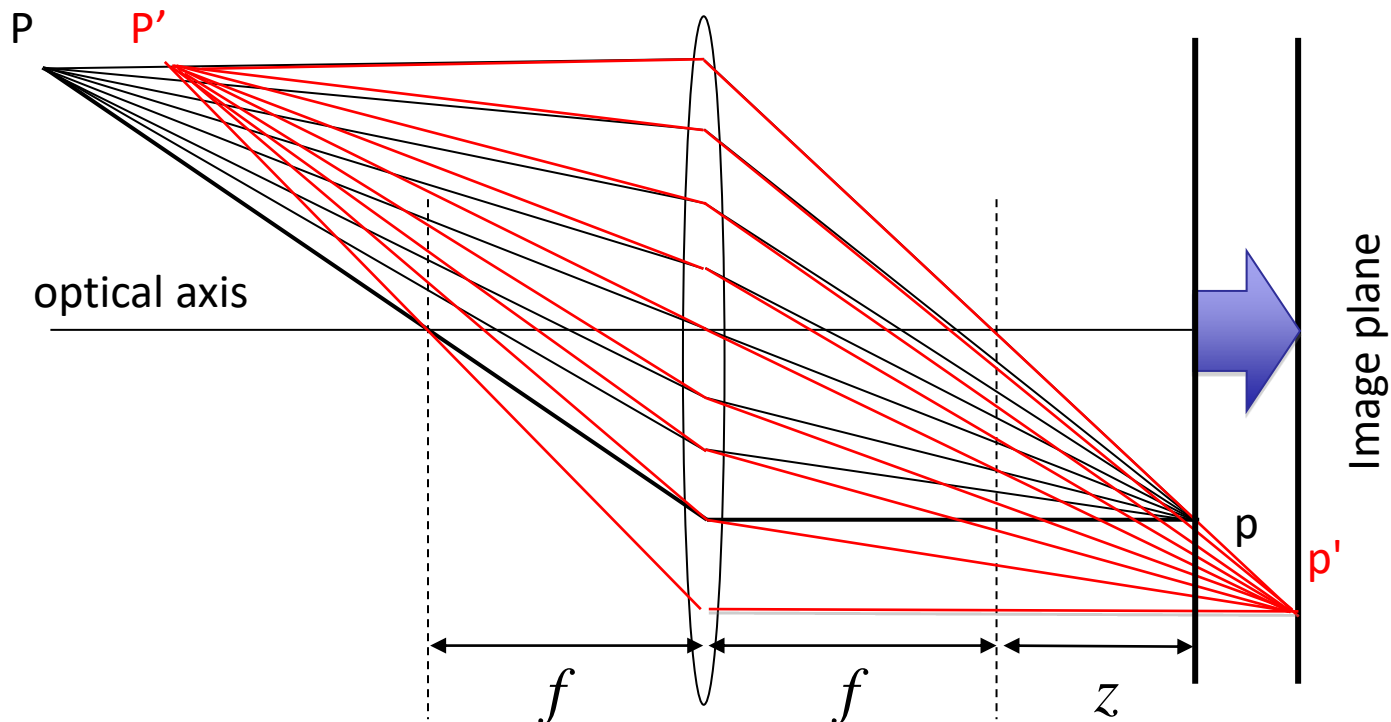
... scales on a lens barrel  
... focal distance oppos  
... are using. If you th  
... the depth of field w  
... e to infinity.◁ For  
... mera has a hyper  
... focus at 18 feet

# Distância focal da lente (3)

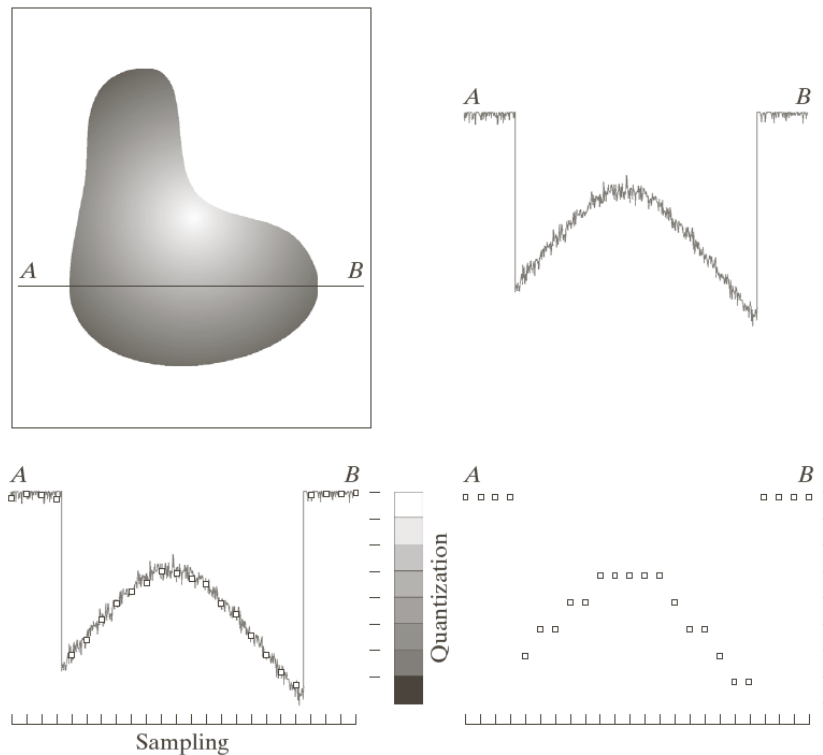


Alteração da profundidade de campo

- Alteração da distância entre a lente e o plano da imagem
- Deformação da lente



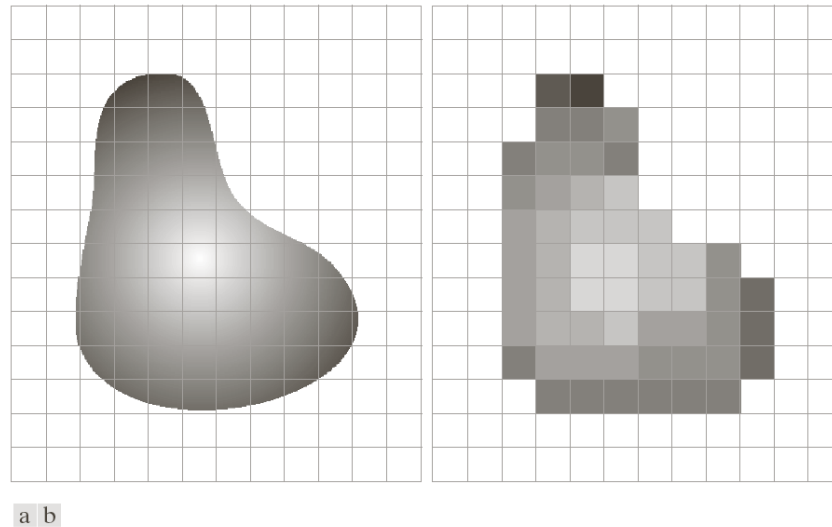
# Amostragem e Quantificação (1)



a b  
c d

**FIGURE 2.16**  
Generating a digital image.  
(a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

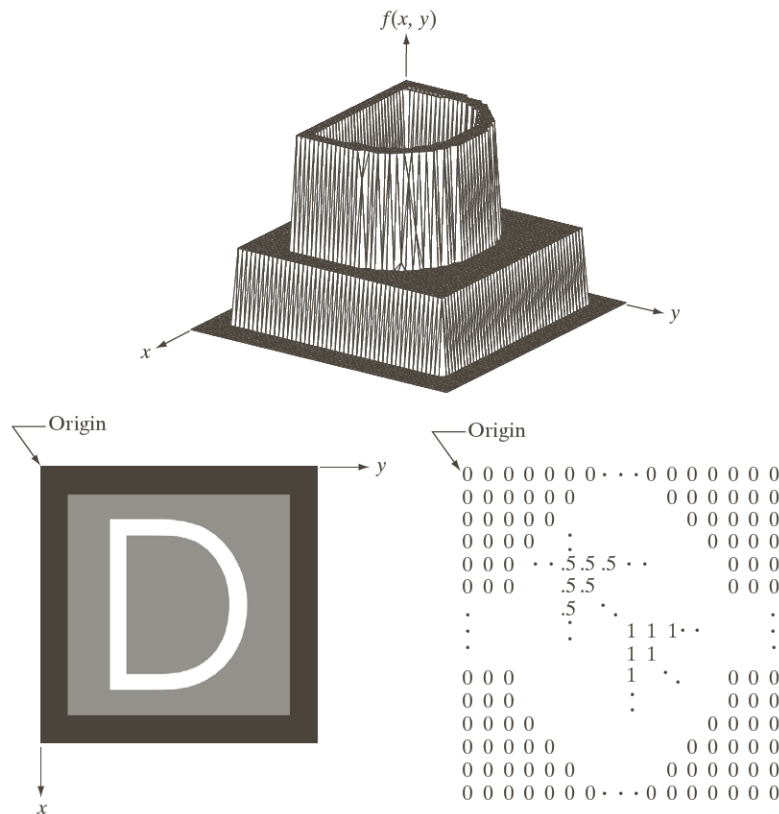
# Amostragem e Quantificação (2)



**FIGURE 2.17** (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.



# Representação de Imagens Digitais (1)



**FIGURE 2.18**  
 (a) Image plotted as a surface.  
 (b) Image displayed as a visual intensity array.  
 (c) Image shown as a 2-D numerical array (0, .5, and 1 represent black, gray, and white, respectively).

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}$$

# Representação de Imagens Digitais (2)



Níveis de Cinzentos ou Número de Cores

$$L = 2^k$$

Número de bits

$$b = M \times N \times k$$

**TABLE 2.1**

Number of storage bits for various values of  $N$  and  $k$ .

$N/k$	1 ( $L = 2$ )	2 ( $L = 4$ )	3 ( $L = 8$ )	4 ( $L = 16$ )	5 ( $L = 32$ )	6 ( $L = 64$ )	7 ( $L = 128$ )	8 ( $L = 256$ )
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912



# Resolução Espacial



a b  
c d

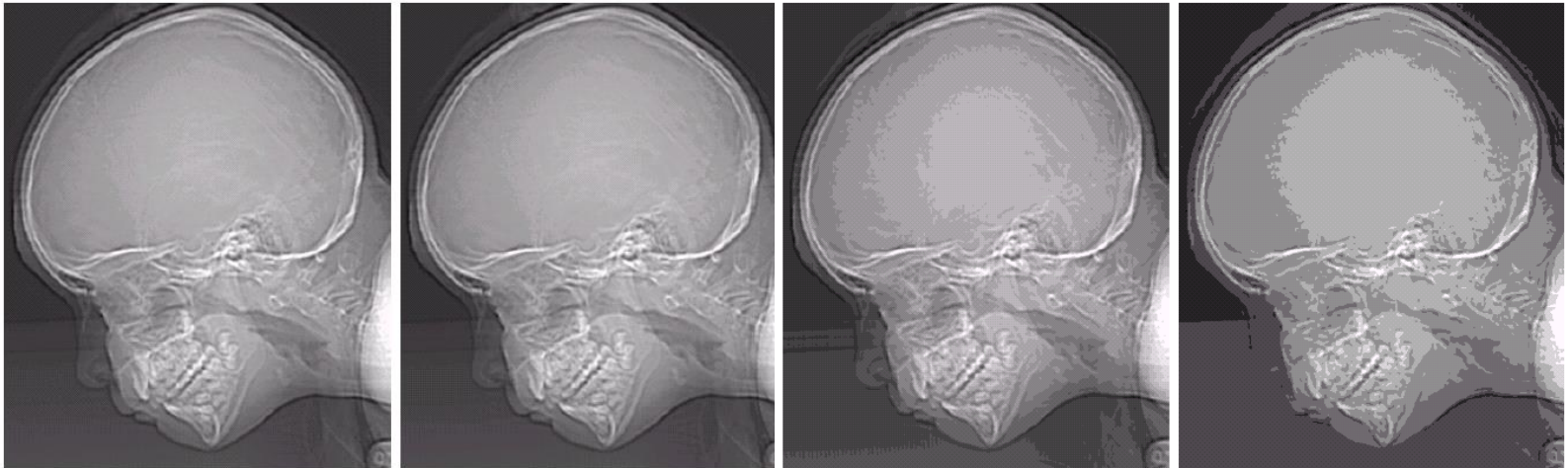
**FIGURE 2.20** Typical effects of reducing spatial resolution. Images shown at: (a) 1250 dpi, (b) 300 dpi, (c) 150 dpi, and (d) 72 dpi. The thin black borders were added for clarity. They are not part of the data.

# Número de Níveis de Quantificação



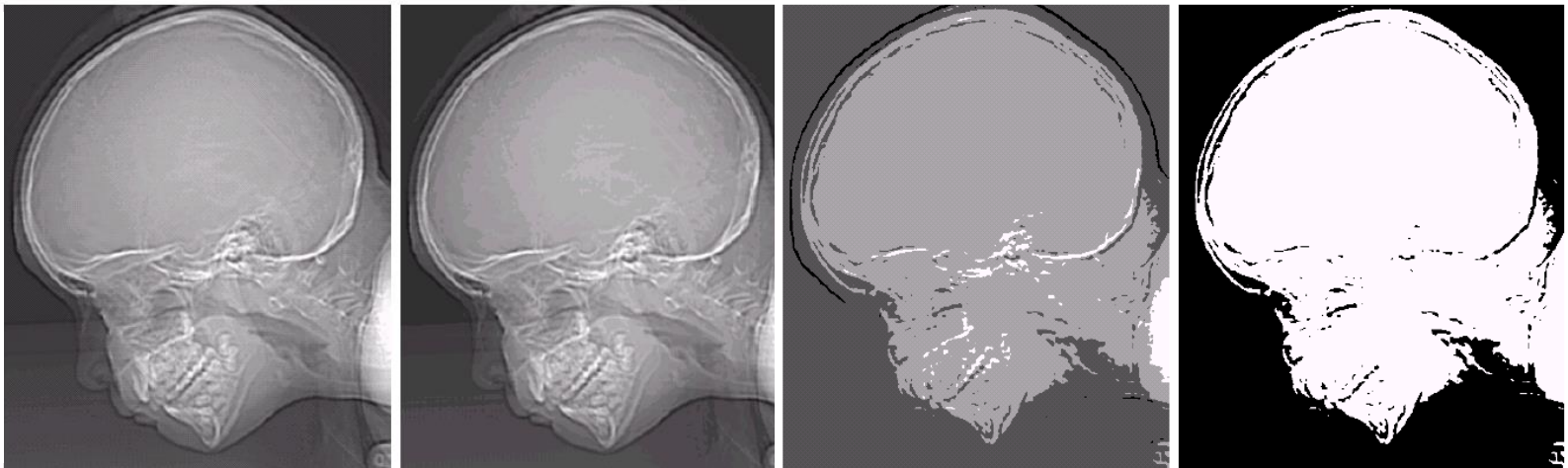
a b  
c d

**FIGURE 2.21**  
(a)  $452 \times 374$ ,  
256-level image.  
(b)–(d) Image  
displayed in 128,  
64, and 32 gray  
levels, while  
keeping the  
spatial resolution  
constant.

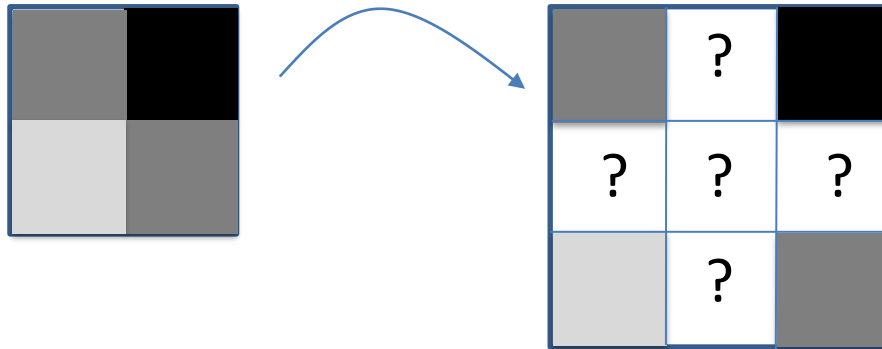


e f  
g h

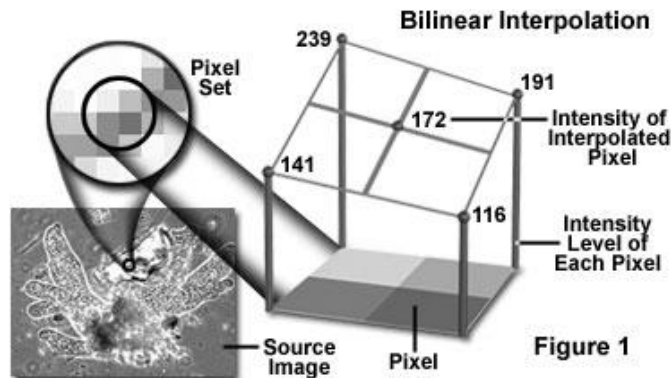
**FIGURE 2.21**  
(Continued)  
(e)–(h) Image  
displayed in 16, 8,  
4, and 2 gray  
levels. (Original  
courtesy of  
Dr. David  
R. Pickens,  
Department of  
Radiology &  
Radiological  
Sciences,  
Vanderbilt  
University  
Medical Center.)



# Interpolação em Imagens (1)

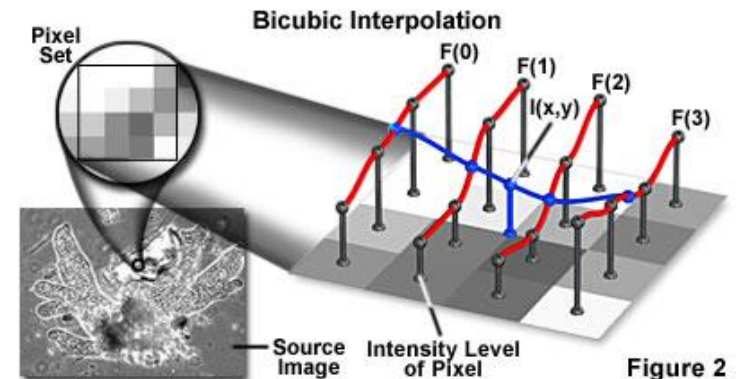


- Vizinho mais próximo
- Interpolação bilinear



$$v(x, y) = ax + by + cxy + d$$

- Interpolação bicúbica



$$v(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j$$

# Interpolação em Imagens (2)



**FIGURE 2.24** (a) Image reduced to 72 dpi and zoomed back to its original size ( $3692 \times 2812$  pixels) using nearest neighbor interpolation. This figure is the same as Fig. 2.20(d). (b) Image shrunk and zoomed using bilinear interpolation. (c) Same as (b) but using bicubic interpolation. (d)–(f) Same sequence, but shrinking down to 150 dpi instead of 72 dpi [Fig. 2.24(d) is the same as Fig. 2.20(c)]. Compare Figs. 2.24(e) and (f), especially the latter, with the original image in Fig. 2.20(a).

# Relações entre Pixéis (1)



## Vizinhança de um pixel

Vizinhança 4  $N_4$

		$y$	
		X	
$x$	X	?	X
		X	

$(x+1, y), (x-1, y), (x, y+1), (x, y-1)$

Vizinhança 8  $N_8$

X	X	X
X	?	X
X	X	X

$(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$



# Relações entre Pixéis (2)



Caminho entre  $p \rightarrow (x, y)$  e  $q \rightarrow (s, t)$

$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$  onde  $(x_0, y_0) = (x, y)$  ,  $(x_n, y_n) = (s, t)$   
e  $(x_i, y_i), (x_{i-1}, y_{i-1})$  são vizinhos, para  $i = 1, 2, \dots, n$

Conectividade

Dois pixéis são conexos se existir um caminho que os une.

Região

Conjunto conexo de pixéis.

1	1	0	1	1	1	0	1
1	1	0	1	0	1	0	1
1	1	1	1	0	0	0	1
0	0	0	0	0	0	0	1
1	1	1	1	0	1	0	1
0	0	0	1	0	1	0	1
1	1	0	1	0	0	0	1
1	1	0	1	0	1	1	1

# Relações entre Pixéis (3)



## Métricas de Distância entre Pixéis

$$p \rightarrow (x, y) \quad q \rightarrow (s, t) \quad z \rightarrow (v, w)$$

$$D(p, q) \geq 0$$

$$D(p, q) = D(q, p)$$

$$D(p, z) \leq D(p, q) + D(q, z)$$

$$\begin{matrix} 2\sqrt{2} & \sqrt{5} & 2 & \sqrt{5} & 2\sqrt{2} \\ \sqrt{5} & \sqrt{2} & 1 & \sqrt{2} & \sqrt{5} \\ 2 & 1 & 0 & 1 & 2 \\ \sqrt{5} & \sqrt{2} & 1 & \sqrt{2} & \sqrt{5} \\ 2\sqrt{2} & \sqrt{5} & 2 & \sqrt{5} & 2\sqrt{2} \end{matrix}$$

Distância Euclidiana  $D_e(p, q) = \sqrt{(x-s)^2 + (y-t)^2}$

$$\begin{matrix} & & & & 2 \\ & & & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \end{matrix}$$

Distância D4 ou *City-block*  $D_4(p, q) = |x-s| + |y-t|$

$$\begin{matrix} & & & & 2 \\ & & & 2 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{matrix}$$

Distância D8 ou *Chessboard*  $D_8(p, q) = \max(|x-s|, |y-t|)$

$$\begin{matrix} & & & & 2 \\ & & & 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{matrix}$$



# Operações Aritméticas (1)



Operações realizadas entre os pares de pixéis correspondentes  
(operação ponta-a-ponto)

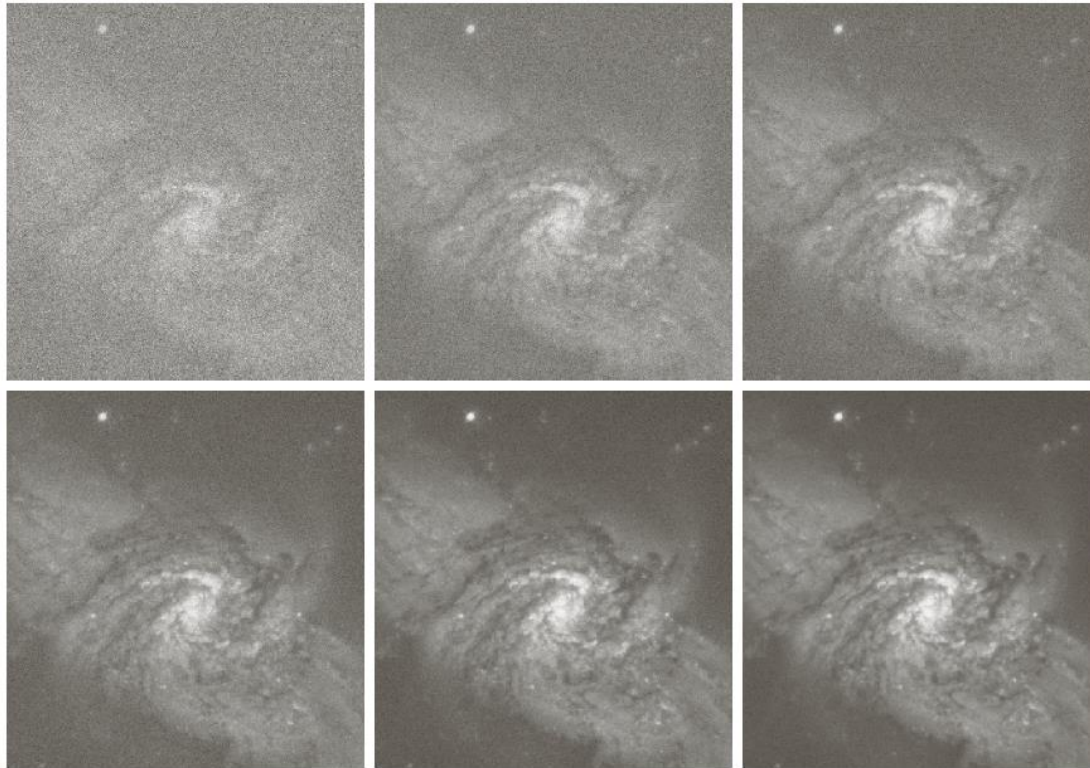
$$s(x, y) = f(x, y) + g(x, y)$$

$$d(x, y) = f(x, y) - g(x, y)$$

$$p(x, y) = f(x, y) \times g(x, y)$$

$$v(x, y) = f(x, y) \div g(x, y)$$

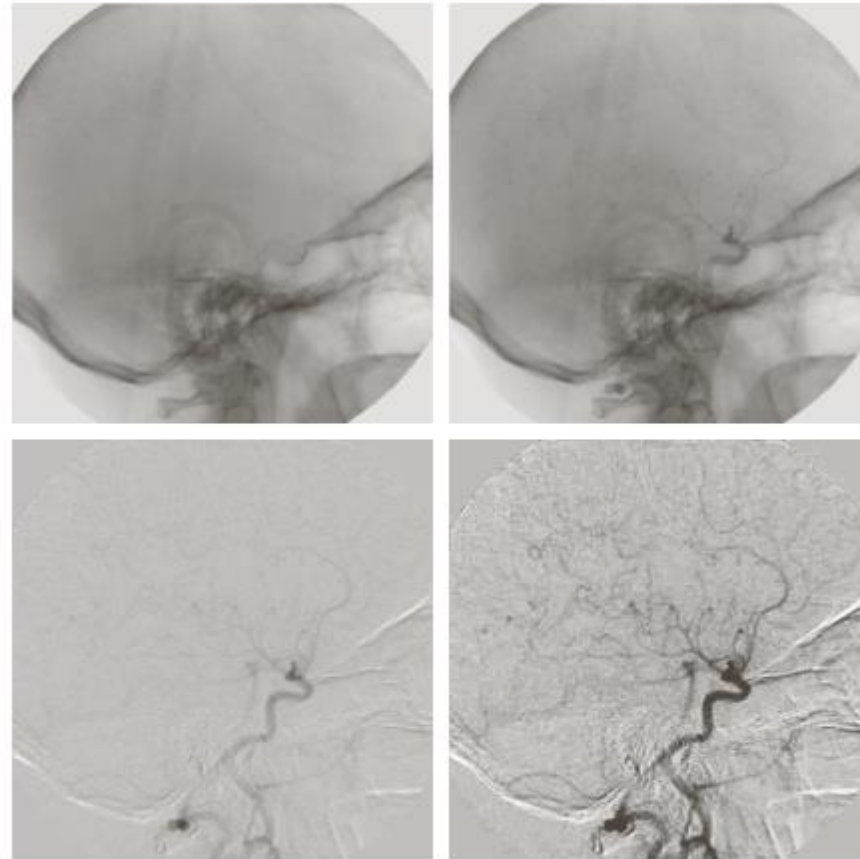
# Operações Aritméticas (2)



a	b	c
d	e	f

**FIGURE 2.26** (a) Image of Galaxy Pair NGC 3314 corrupted by additive Gaussian noise. (b)–(f) Results of averaging 5, 10, 20, 50, and 100 noisy images, respectively. (Original image courtesy of NASA.)

# Operações Aritméticas (3)



a b  
c d

**FIGURE 2.28**

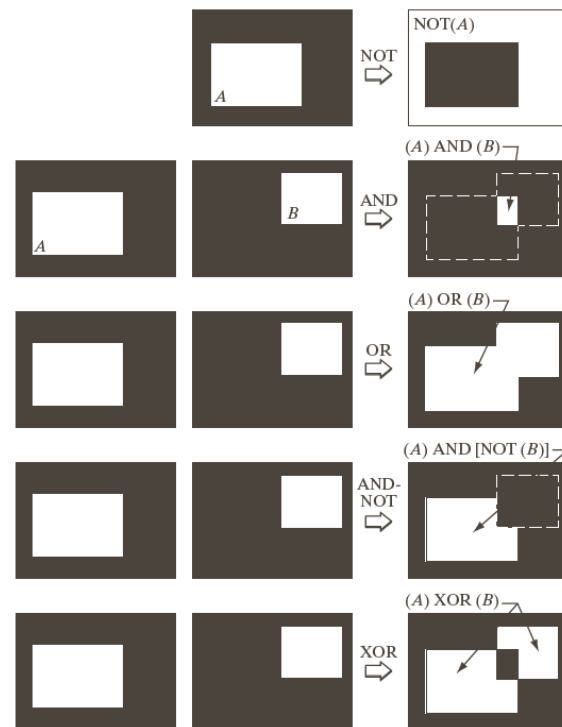
Digital subtraction angiography. (a) Mask image. (b) A live image. (c) Difference between (a) and (b). (d) Enhanced difference image. (Figures (a) and (b) courtesy of The Image Sciences Institute, University Medical Center, Utrecht, The Netherlands.)

# Operações Aritméticas (4)



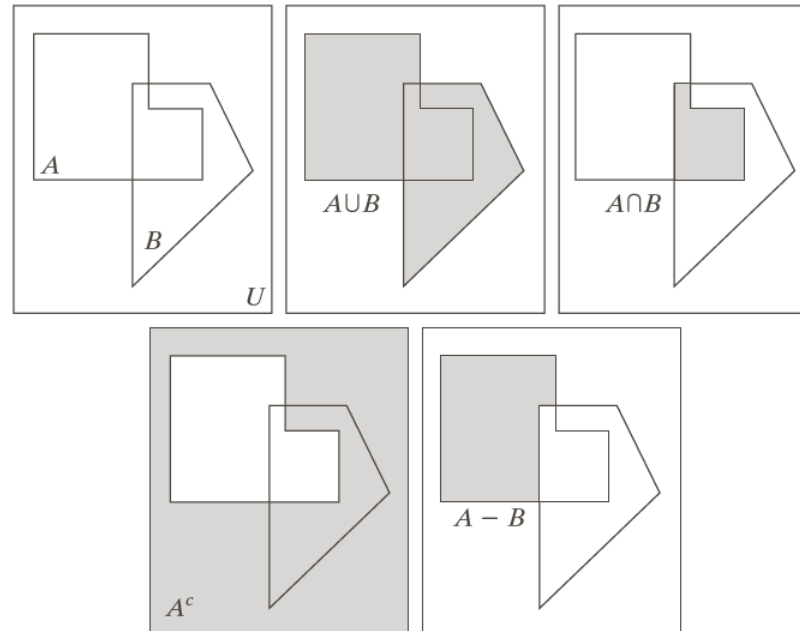
**FIGURE 2.30** (a) Digital dental X-ray image. (b) ROI mask for isolating teeth with fillings (white corresponds to 1 and black corresponds to 0). (c) Product of (a) and (b).

# Operações Lógicas e de Conjuntos (1)



**FIGURE 2.33**  
Illustration of logical operations involving foreground (white) pixels. Black represents binary 0s and white binary 1s. The dashed lines are shown for reference only. They are not part of the result.

# Operações Lógicas e de Conjuntos (2)



a	b	c
d	e	

**FIGURE 2.31**

(a) Two sets of coordinates,  $A$  and  $B$ , in 2-D space. (b) The union of  $A$  and  $B$ . (c) The intersection of  $A$  and  $B$ . (d) The complement of  $A$ . (e) The difference between  $A$  and  $B$ . In (b)–(e) the shaded areas represent the member of the set operation indicated.

# Operações Espaciais (1)



1. Operação com base num pixel

$$s = T(z)$$

2. Operação com base em vizinhos

$$S_{xy} \text{ vizinhos do pixel } p \rightarrow (x, y)$$

3. Transformações geométricas

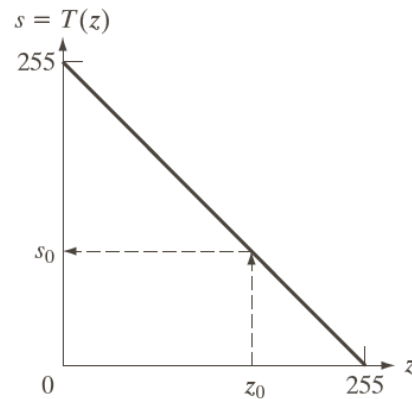
$$(x, y) = T\{(v, w)\}$$



# Operações Espaciais – Exemplos (1)



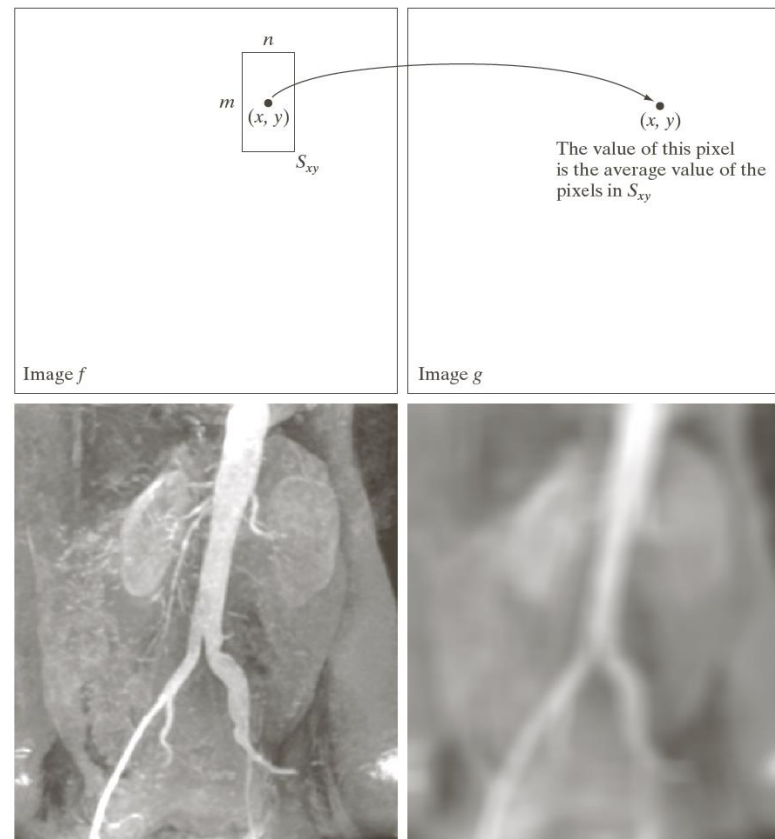
## Imagem Negativa



**FIGURE 2.34** Intensity transformation function used to obtain the negative of an 8-bit image. The dashed arrows show transformation of an arbitrary input intensity value  $z_0$  into its corresponding output value  $s_0$ .

# Operações Espaciais – Exemplos (2)

## Média



a b  
c d

**FIGURE 2.35**  
Local averaging using neighborhood processing. The procedure is illustrated in (a) and (b) for a rectangular neighborhood. (c) The aortic angiogram discussed in Section 1.3.2. (d) The result of using Eq. (2.6-21) with  $m = n = 41$ . The images are of size  $790 \times 686$  pixels.

# Operações Espaciais – Exemplos (3)



## Transformações Afim

$$[x \ y \ 1] = [v \ w \ 1]T = [v \ w \ 1] \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{bmatrix}$$

**TABLE 2.2**

Affine transformations based on Eq. (2.6.–23).

Transformation Name	Affine Matrix, T	Coordinate Equations	Example
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v$ $y = w$	
Scaling	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = c_x v$ $y = c_y w$	
Rotation	$\begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v \cos \theta - w \sin \theta$ $y = v \sin \theta + w \cos \theta$	
Translation	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix}$	$x = v + t_x$ $y = w + t_y$	
Shear (vertical)	$\begin{bmatrix} 1 & 0 & 0 \\ s_v & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v + s_v w$ $y = w$	
Shear (horizontal)	$\begin{bmatrix} 1 & s_h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v$ $y = s_h v + w$	

# Operações Espaciais – Exemplos (4)



Mapeamento directo

$$(x, y) = T\{(v, w)\}$$

Mapeamento inverso

$$(v, w) = T^{-1}\{(x, y)\}$$



a b c d

**FIGURE 2.36** (a) A 300 dpi image of the letter T. (b) Image rotated 21° clockwise using nearest neighbor interpolation to assign intensity values to the spatially transformed pixels. (c) Image rotated 21° using bilinear interpolation. (d) Image rotated 21° using bicubic interpolation. The enlarged sections show edge detail for the three interpolation approaches.

# Transformadas



Domínio espacial

Domínio da transformada

$$f(x, y)$$

$$T(u, v)$$

$$T(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) r(x, y, u, v)$$

Separável

$$r(x, y, u, v) = r_1(x, u) r_2(y, v)$$

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u, v) s(x, y, u, v)$$

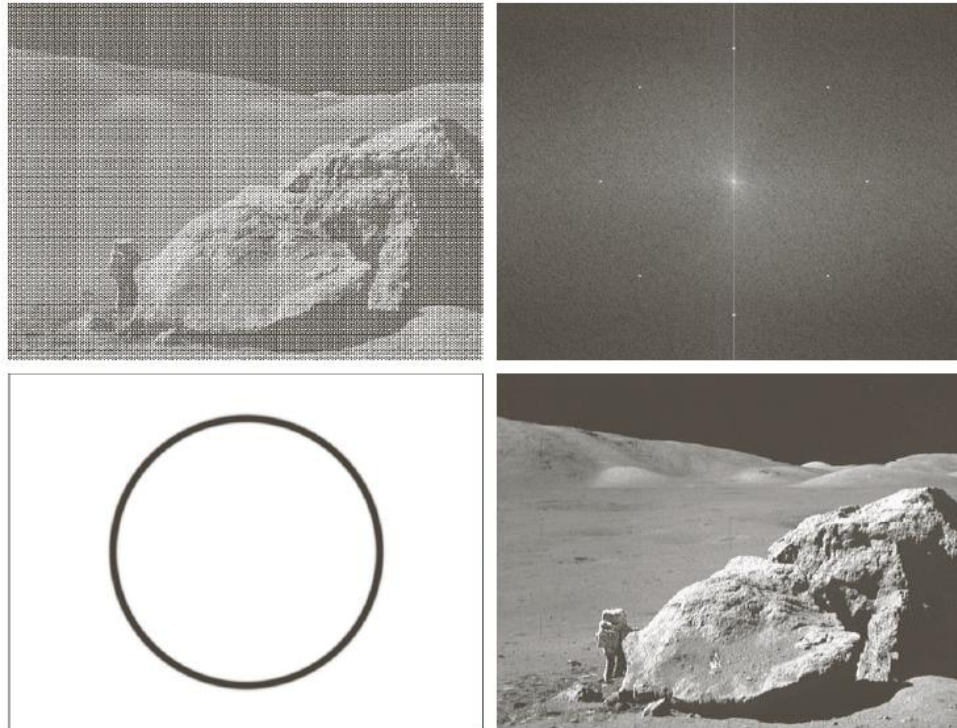
Simétrica

$$r(x, y, u, v) = r_1(x, u) r_1(y, v)$$



**FIGURE 2.39**  
General approach  
for operating in  
the linear  
transform  
domain.

# Transformadas -Exemplo



a b  
c d

**FIGURE 2.40**

(a) Image corrupted by sinusoidal interference. (b) Magnitude of the Fourier transform showing the bursts of energy responsible for the interference. (c) Mask used to eliminate the energy bursts. (d) Result of computing the inverse of the modified Fourier transform. (Original image courtesy of NASA.)

# Tipos de Imagens



- **Definição:** uma imagem monocromática (cinzentos),  $I[r,c]$ , associa a cada pixel um valor escalar (intensidade)
- **Definição:** uma imagem multiespectral,  $M[r,c]$ , associa a cada pixel um vector de dimensão  $N$ . No caso de imagens a cores tem-se  $N=3$  (RGB)
- **Definição:** uma imagem binária,  $B[r,c]$ , associa a cada pixel apenas os valores 0 ou 1
- **Definição:** uma imagem de classes (*labels*),  $L[r,c]$ , associa a cada pixel um dado símbolo, escolhido de entre os símbolos de um alfabeto finito



# Formatos de imagens

- Cabeçalho (*header*)
- Dados
  - Utilização (ou não) de algoritmos de compressão de dados.
    - Compressão sem perdas (*lossless*) ou com perdas (*lossy*)



Exemplo de formato →

```
P2
# sample small picture 8 rows of 16 columns, max grey value of 192
# making an image of the word "Hi".
16 8 192

64 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64
64 64 128 128 64 64 64 128 128 64 64 192 192 64 64 64
64 64 128 128 64 64 64 128 128 64 64 192 192 64 64 64
64 64 128 128 128 128 128 128 128 64 64 64 64 64 64 64
64 64 128 128 128 128 128 128 128 64 64 128 128 64 64 64
64 64 128 128 64 64 64 128 128 64 64 128 128 64 64 64
64 64 128 128 64 64 64 128 128 64 64 128 128 64 64 64
64 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64
```



Image File Format	No. Bytes "Hi"	No. Bytes "Cars"
PGM	595	509,123
GIF	192	138,267
TIF	918	171,430
PS	1591	345,387
HIPS	700	160,783
JPG (lossless)	684	49,160
JPG (lossy)	619	29,500

# Leitura Sugerida



- Capítulos 1 e 2 de R. Gonzalez, R. Woods, “Digital Image Processing”, 3ª edição, 2008.
- Capítulo 2 de L. Shapiro, G. Stockman, “Computer Vision”, 2001.